

**UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF NEW JERSEY**

<p>PEARSON EDUCATION, INC.,</p> <p style="text-align: center;">Plaintiff,</p> <p style="text-align: center;">v.</p> <p>CHEGG, INC.,</p> <p style="text-align: center;">Defendant.</p>	<p>Civil Action No.: 21-16866</p> <p style="text-align: center;">OPINION AND ORDER OF THE SPECIAL DISCOVERY MASTER RE: READABLE TBS DATA</p>
---	--

LINARES, J.

This matter comes before the Special Master by way of Plaintiff Pearson Educations, Inc.’s July 13, 2023 letter application (“Motion”) seeking an Order compelling Chegg to comply with the Special Master’s June 5, 2023 Order requiring Defendant Chegg, Inc. to produce certain TBS data in a “readable” format. (ECF No. 135). Chegg submitted its opposition to the application on July 20, 2023 (“Chegg Opp.”)¹, to which Pearson replied. (“Pearson Rep.”). For the reasons set forth below, Pearson’s letter application is denied.

I. BACKGROUND

The Special Masters presumes the parties familiarity with the factual background and procedural posture of this matter. Accordingly, the Special Master will only recite the facts relevant to the disposition of the subject dispute.

¹ In its opposition, Chegg claims that Pearson’s letter application does not contain a clear request for relief. (Chegg Opp. n.1). While Chegg is correct that the Pearson application does fail to state a succinct request for relief, it is apparent to the Special Master that Pearson seeks compliance with the pertinent portion of ECF No. 135.

One of the issues raised by Pearson prior to the April 4, 2023 hearing before the Special Master involved the “readability” of certain data contained in Chegg’s TBS Database. (*See generally* Motion, Chegg Opp.). During the hearing, the Parties explained that the TBS Database housed all of Chegg’s solutions for the end-of-chapter questions that form the basis of the underlying copyright dispute. (April 4, 2023 Hearing Transcript (“Hrg. Tr.”) 57:24-58:5, 158:19-22).

During the course of discovery, both before and after the Special Master’s appointment, Pearson has sought to obtain information from the TBS Database. (*See, e.g.*, Hrg. Tr. 45:17-19 (Pearson explaining the various Requests for Production that cover demands for information and data from the TBS Database)). Pearson, in part, asserted that the data Chegg had previously produced from the TBS Database was “unreadable,” because it contained “rogue” alphanumeric characters. (Hrg. Tr. 52:21-23 (“MR. OPPENHEIM: It’s not even close to what we’ve asked for. First of all, the questions they provided is (*sic*) unreadable.”)).

In its post-hearing submission on this topic, Chegg agreed to produce the data in a different format that it anticipated would be easier for Pearson to understand. (ECF No. 135 at 6). Accordingly, the June 5, 2023 Opinion and Order memorialized Chegg’s agreement and required Chegg to “produce the data with respect to the text of Pearson’s questions [from the TBS Database] in a ‘readable’ format that will be easier for Pearson to understand.”). (ECF No. 135 at 8) .

In the present letter application, Pearson maintains that Chegg still has not produced the demanded discovery in a “readable format.” (Motion at 1 (internal quotations omitted)). Specifically, Pearson explains that Chegg’s most recent production still contains “coded markups that were not part of the question text (*e.g.*, “<” and “>”).” (Id.). Pearson further avers that

Chegg's production also contains "substantive characters in an algebraic equation." (Pearson Rep. at 1).

Chegg's opposition begins by outlining the tortured history of this dispute. (Chegg Opp. at 1-2). In that long explanation, Chegg notes that the characters that Pearson asserts render the data "unreadable" are a product of the various computer languages used to store the data and the export formats used to retrieve same. (Id.). Indeed, Chegg's first production was done in comma separated value or ".csv" format, which, as its name suggests, uses commas to separate the various pieces of data. (Id. at 1).

Before the Special Master could make any specific findings regarding the validity or appropriateness of the .csv format, Chegg offered to re-produce the data in a different format. As a result, the Special Master issued the June 5, 2023 Opinion and Order, which, in part, required Chegg to produce the data in a more user friendly or "readable" form, consistent with its on-the-record representations. (ECF No. 135 at 8).

Since then, Chegg re-produced the data using a different file format; ".json." (Chegg Opp. at 2). Pearson acknowledges that Chegg has re-produced the data in the .json file format, but continues to take issue with certain "problematic markups." (Motion at 2). This letter application followed after the Parties' meet-and-confer efforts on the issue failed.

II. ANALYSIS

The plain language of Rule 34 alone disposes of the dispute. Specifically, subsection (b)(2)(E)(i) is instructive and states that "A party must produce documents as they are kept in the usual course of business or must organize and label them to correspond to the categories in the

request.” Fed. R. Civ. P. 34(b)(2)(E)(i). Moreover, as to electronically stored information (“ESI”) “a party must produce it in a form or forms in which it is ordinarily maintained.” (Id. at (ii)). The Third Circuit has found that no additional manipulation or explanation is required when a party does in fact produce the demanded information “as [it is] kept in the normal course of business.” *Le v. City of Wilmington*, 480 F. App’x 678, 687 (3d Cir. 2012).

Here, Chegg has produced the data in several forms. Chegg’s first attempt was done in the above described .csv format. The use of commas in this type of file is inherent and necessary for the export. Nonetheless, Pearson asserted that the data as it was presented in the .csv file format was unusable, due to all of the seemingly random commas associated with this file format.

In an attempt to alleviate Pearson’s concerns, and to comply with the Special Master’s subsequent Order, Chegg re-produced the data in the .json format. According to Chegg, this removed many of the unnecessary characters. But, as a result of how Chegg stored or inputted the information into the database, source code or program specific code is still present in the .json reproduction. Pearson continued to object to the production, categorizing it as “unreadable.”

Based on the foregoing, the Special Master concludes that Chegg is in compliance with the June 5, 2023 Opinion and Order and its discovery obligations under the Federal Rules. Indeed, Chegg has represented that the data contained in the .json file is how Chegg maintains same on the TBS Database in the ordinary course of business. That is the beginning and the end of the inquiry. This is the only burden Chegg has in connection with producing the demanded ESI. *See Le*, 480 F. App’x at 687.

The *Le* case is instructive. There, the plaintiff was dissatisfied with defendant’s document production because it did not have labels nor were the documents categorized in a way that allowed

the plaintiff to easily ascertain which discovery demands the electronic discovery was responsive to. *Id.* at 686. The District Court denied the plaintiff’s motion to compel citing Rule 34(b)(2)(E). *Id.* The Third Circuit agreed with the District Court’s ruling and explained that “such labeling is not required when the party otherwise complies with the rule by producing the documents *as they are kept in the normal course of business.*” *Id.* at 686-87 (emphasis added).

Similarly here, Pearson is looking to obtain a “clean” copy of the data in the TBS Database. However, as Chegg has represented, that precise version of the data does not exist. Chegg has satisfied its discovery obligations by producing the data – in two different formats – as Chegg maintains it in the ordinary course of business. Any Order requiring further manipulation of the data would be improper as it would ask Chegg to exceed its obligations under the Federal Rules.

Anecdotally, the Special Master notes that the excerpts of data presented by Pearson do not appear to be infused with a significant amount of unknown characters such that they become completely unreadable or useless. Based on what was presented to the Special Master in Pearson’s letter, along with the fact that Chegg has provided a proverbial legend for the random characters, the Special Master concludes that Chegg has in fact provided Pearson with a more readable version of the TBS Database.

III. ORDER

IT IS on this 9th day of August, 2023,

ORDERED that Pearson's letter application an Order compelling Chegg to comply with ECF No. 135 by providing "readable" TBS data is hereby **DENIED**.

SO ORDERED.

/s/ Jose L. Linares

Hon. Jose L. Linares, U.S.D.J. (Ret.)